Welcome to the Benchmark Dose models training session. This is the 3rd module and focuses on modeling dichotomous cancer data.

2. Here is a list of the available models for analyzing dichotomous data in BMDS. The multistage model has a long history of use for tumor dose-response analysis and has been modified to make it more suitable for cancer risk assessment per EPA's 2005 cancer guidelines It will be the focus of this training module. However, the EPA 2005 cancer guidelines do not discourage the use of other models to obtain a cancer point of departure (POD).

3. The form of the multistage model is shown here. There are two terms in this model. The gamma term is a background term and the beta term represents the power. The current guidance suggests that you restrict the betas to be greater than or equal to zero. If the betas take on values less than zero, the dose-response curve can become wavy.

4. The differences between the multistage-cancer model and the multistage model are listed on this slide. In the multistage-cancer model, beta is always restricted so this will no longer be an option on the model run page. The cancer slope factor is calculated automatically and shown on the text output and a linear extrapolation appears on the plot or graphical output. Unlike the other BMDS dichotomous models, both of the BMDS multistage models present a BMDU (an estimate of the 95% upper confidence limit on the BMD).

5. An example of the graphical output from the multistage-cancer model is shown. Notice that the blue line is a linear extrapolation which provides the point of departure for this data.

6. The cancer slope factor or CSF is calculated and will appear on the text output. The CSF is the benchmark response level divided by the lower limit on the benchmark dose at that same response level. The CSF can be obtained from the BMDL output of other models, but must be manually calculated per this simple equation.

7. The graphical output provided on this slide shows where these features will be visible. The cancer slope factor appears as a linear extrapolation, and the point of departure for this data set and BMR level is provided.

8. Now we will look at an example of using the multistage-cancer model.

9. The following data set is the one we will use for running the multistage-cancer model. There are 4 dose groups in addition to the control and there appears to be a dose-response relationship.

10. The BMD Analysis flowchart shown on this slide is designed to walk the user through the BMD analysis steps. First, we will choose the BMR's to evaluate.

11. The guidance on selecting a BMR for cancer data says that an extra risk of 10% is the default since the 10% response is at or near the limit of sensitivity in most cancer bioassays and in some non-cancer bioassays.
There are examples in risk assessment where one uses a lower BMR, typically if a study has greater than usual sensitivity. The BMD and BMDL 10 should always be presented for comparison purposes

12. The BMR can be set for the cancer model on the model run screen. The green arrow points to the box where the BMR is entered.

13. After choosing a BMR to evaluate, we must next, select a model.

14. On the Create/Edit dataset screen, we will select dichotomous as the model type and select multistage-cancer as the model.

15. Next we will set the parameters.

16. The only change we will need to make is to the degree of the polynomial. This should be set to 1. The risk type is set to extra, the BMR is at 0.1 or 10%, the confidence level is 0.95 or 95% and the BMD calculation box is checked.

17. Now we will run the model.

18. Several screens will pop up following running the model. One is the graphical output which is shown here. This shows you the visual fit of the model to the data. Here we can see how the multistage-cancer model fits all 5 data points.

19. The next screen you will see is the text output screen, a portion of which is shown on this slide. In running this data, there are certain values that we will record for comparison later. The red arrows point to the main values you will record. These include, the p value, the BMD and BMDL and the cancer slope factor. We will also examine the scaled residuals which are found in the goodness of fit table.

20. Our next decision to make is to determine does the model fit the data.

21. To determine this, there are several tools you can use. First, a global measurement of fit is found in the goodness of fit p value. A value of p greater than 0.1 indicates that the model fits the data.

22. This global measurement is based upon a chi square calculation. The equation and the definition of the terms are shown.

23. The p value can be found on the text output file under the goodness of fit table. The p-value is shown here circled in green.

24. The next tool we can use to determine if the model fits the data are the local measurements of fit called the scaled residuals. Scaled residuals should have an absolute value less than 2.0.

25. How do we use the scaled residuals to assess how well the model fits the low dose fit. Occasionally the software will converge on a less than optimal solution or will "fit' the wrong region of the dose response curve, the high end for example. Scaled residuals are a measure of how closely the model fits the data at each point. They are calculated by taking the observed data point and subtracting the estimated data point, with 0 equaling an exact fit. Absolute values near the BMR should be lowest and the technical guidance says you should question values of scaled residuals that are greater than an absolute value of 2.

26. These values appear on the text output in the column marked scaled residuals under the goodness of fit table. We are particularly interested in the scaled residual closest to the BMR which in this case is the dose of 50. The absolute value is greater than 2, so we should question whether or not the model is fitting the data well as suggested by our guidance.

27. Looking at our graph of the multistage-cancer model fit to our data, the scaled residuals look at how close the model or red line come to the data points, shown in green.

28. The next step in our flowchart is to consider have all the models/model options been considered?

29. Here, we will continue to run the multistage-cancer but we will change the degree of the polynomial to see if we can achieve a better fit. On the Type Model run slide shown, the green arrow points to the degree of the polynomial box. We will now change this from a 1 to a 2 and rerun the model.

30. This slide shows the graph for the fit of the 2nd degree multistage-cancer model to all 5 data points. Visually, there appears to be an improvement in fit from the first degree model.

31. This slide shows the graph for the fit of the 3rd degree multistage-cancer model to all 5 data points. There does not appear to be an improvement in the fit from the 2nd degree model.

32. Here we show why there is not an improvement in the fit from the 2nd to the 3rd degree cancer model. Under the parameter estimates table on the text output, we can see that even when the degree of the polynomial is set to 3, the model with the best fit to the data uses beta 2, but not beta 1 or beta 3. Thus, it results in an identical model fit to that obtained when the degree of the polynomial was set to 2.

33. Here we have provided a summary of the cancer model fit to all five data points. We have recorded the BMD10, the BMDL 10, the AIC the p value and the scaled residuals for the 1st, 2nd and3rd degree multistage-cancer models.

34. To summarize this section, we can make the following observations. First, the multistage-cancer model does not provide a satisfactory fit to the data that is, we do not have a p value greater than 0.1. The poor model fit may be due to the model trying to fit the high dose data as shown by the scaled residuals. This results in a poor fit to the low dose region which is the region we are most interested in. We will consider removing the highest dose group.

35. We will now rerun the multistage-cancer model after removing the highest dose group. Right away we can see that the visual fit has dramatically improved the visual fit.

36. Therefore, removing the highest dose group results in a good fit to the data.

37. Here we have provided a summary of the cancer model fit to the four data points similar to that presented on slide 35. We can see that the 3rd degree multistage-cancer model has a very good p value of 0.9709 and the scaled residuals have improved, especially near the 50 mg/kg dose which is near the BMR.

38. To summarize this section, based on the statistical and visual comparison, removing the highest dose will result in a better model fit to the data. The AIC and scaled residuals can be used to assess how well the model fits a data point. Visual inspection can be used to determine whether it is appropriate to model all of the data.

39. Continuing our summary, the 3rd degree cancer model provides the best data fit. It had the highest p value, the smallest AIC and the smallest scaled residuals at the data point close to the BMR as compared to the 1st and 2nd degree cancer model. The 4th degree model provides identical model estimates and therefore, we should use the simpler model. Thus, the results from the 3rd degree model fit will be used for the cancer risk assessment.

40. Finally, according to our flowchart, we will use the BMDL from the model that provides the best fit and document the BMD analysis as outlined in reporting requirements of the Benchmark Dose Technical Guidance document.

41. The text output from the 3rd degree multistage-cancer model is shown and provides us with the information needed for the last step.

42. The cancer slope factor as we stated earlier, is calculated for you and appears on the text output for the multistage-cancer model. It is calculated by dividing the BMR here 0.1, by the BMDL at that BMR, which is 52.28 in this case, to arrive at the cancer slope factor, 0.001913. Note that because of the slope of the dose-response curve, lower BMRs would result in lower CSF and unit risk estimates.